Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **479,** 1807–1821 (2018)
Advance Access publication 2018 May 31

doi:10.1093/mnras/sty1424

The gas-phase metallicities of star-forming galaxies in aperture-matched SDSS samples follow potential rather than mass or average surface density

Francesco D'Eugenio,[1,2,3]★ Matthew Colless,[1,2,4] Brent Groves,[1,4] Fuyan Bian[1]† and Tania M. Barone[1,2,4,5]

[1]*Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia*
[2]*Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO)*
[3]*Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium*
[4]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*
[5]*Sydney Institute for Astronomy, School of Physics, The University of Sydney, NSW 2006, Australia*

Accepted 2018 May 30. Received 2018 May 29; in original form 2018 January 31

ABSTRACT

We present a comparative study of the relation between the aperture-based gas-phase metallicity and three structural parameters of star-forming galaxies: mass ($M \equiv M_*$), average potential ($\Phi \equiv M_*/R_e$), and average surface mass density ($\Sigma \equiv M_*/R_e^2$, where R_e is the effective radius). We use a volume-limited sample drawn from the publicly available SDSS DR7, and base our analysis on aperture-matched sampling by selecting sets of galaxies where the SDSS fibre probes a fixed fraction of R_e. We find that between 0.5 and 1.5 R_e, the gas-phase metallicity correlates more tightly with Φ than with either M or Σ, in that for all aperture-matched samples, the potential–metallicity relation has (i) less scatter, (ii) higher Spearman rank correlation coefficient, and (iii) less residual trend with R_e than either the mass–metallicity relation and the average surface density–metallicity relation. Our result is broadly consistent with the current models of gas enrichment and metal loss. However, a more natural explanation for our findings is a local relation between the gas-phase metallicity and the escape velocity.

Key words: galaxies: abundances – galaxies: evolution – galaxies: fundamental parameters.

1 INTRODUCTION

The abundances of elements heavier than helium in stars and in the interstellar medium (metallicity) evolve over cosmic time alongside the host galaxies. The link between chemical evolution and galaxy formation is engraved in the empirical correlations between metallicity and other physical properties of galaxies, such as total mass (Lequeux et al. 1979; Garnett & Shields 1987; Vila-Costas & Edmunds 1992).

The advent of large, single-fibre spectroscopic surveys allowed the study of the systematic variation of metallicity for statistically significant samples (Sloan Digital Sky Survey [SDSS], York et al. 2000; 6dF Galaxy Survey, Jones et al. 2004). For star-forming galaxies, the gas-phase metallicity averaged over a fibre aperture was found to correlate tightly with stellar mass (Tremonti et al. 2004; hereafter T04). These authors determined the gas-phase metallicity using strong emission lines in a sample of galaxies drawn from the SDSS Data Release 2 (Abazajian et al. 2004). For passive galaxies, it is typically easier to measure the metallicity of the stars, usually by comparing a set of absorption line indices (Lick index system; Worthey et al. 1994) to a grid of models (e.g. Worthey 1994; Thomas et al. 2010). The stellar metallicity measured within the SDSS fibres correlates with both the stellar mass of the galaxy and the root mean square velocity of the line-of-sight velocity distribution (σ; Thomas et al. 2010).

Alongside these aperture-averaged relations, metallicity is known to vary systematically within each galaxy: the existence of radial metallicity gradients has been known for more than two decades, for both the gas-phase metallicity (e.g. Zaritsky, Kennicutt & Huchra 1994) and the stellar metallicity (e.g. Davies, Sadler & Peletier 1993). More recently, we have been able to study the resolved metallicity relations with unprecedented detail and with much larger samples of galaxies, thanks to new large integral-field spectroscopy surveys (IFS; SAURON, de Zeeuw et al. 2002, ATLAS^{3D}, Cappellari et al. 2011, CALIFA, Sánchez et al. 2012, SAMI, Croom et al. 2012 and MaNGA, Bundy et al. 2015). For the gas-phase metallicity, the aperture-averaged mass–metallicity relation is mirrored by two local relations: the existence of a universal radial gradient (Sánchez et al. 2014) and the relation between local metallicity and local stellar mass surface density (Moran et al. 2012; Rosales-Ortega et al. 2012; Sánchez et al. 2013; Barrera-Ballesteros et al. 2016). The relation with stellar mass surface density appears tighter than

★ E-mail: francesco.deugenio@ugent.be
† Stromlo Fellow.

the physically driven relations with gas fraction or escape velocity (V_{esc}; Barrera-Ballesteros et al. 2018; hereafter BB18).

For stellar metallicity, the best aperture-averaged predictor is related to the gravitational potential, rather than either stellar mass or surface mass density, as determined using a comparative analysis (Barone et al. 2018; hereafter B18). Li et al. (2018) also find that stellar metallicity is approximately constant along lines of constant σ, in agreement with the quantitative analysis of B18. Although there is no comparable study of the resolved relations for stellar metallicity, the Mgb absorption index follows a tight relation with V_{esc} (Emsellem et al. 1996). This relation is both global and local, i.e. the trend between the local Mgb and V_{esc} is the same across all galaxies, as well as for the values of Mgb and V_{esc} measured inside one effective radius (R_e; Scott et al. 2009).

Determining whether stellar metallicity and gas-phase metallicity follow different physical properties of galaxies (and if so, why) is crucial to our understanding of how galaxies form and evolve. The fact that we cannot state a priori what is the best predictor of gas-phase metallicity demonstrates that the physical origin of the metallicity relations is still not fully understood. Even less clear is why stellar metallicity correlates with V_{esc}, whereas previous studies of the metallicity of star-forming gas propose primary correlations with stellar mass (or surface mass density).

In order to address this problem, we conduct the first comparative analysis of the relation between gas-phase metallicity and three structural properties of galaxies: mass (M), gravitational potential (Φ), and surface mass density (Σ). We use a volume-limited sample of single-fibre spectroscopy data drawn from SDSS; in order to overcome aperture bias inherent to single-fibre data, we introduce the concept of aperture-matched subsampling: considering subsets of galaxies with a given physical size (in units of R_e) equal to the aperture of the SDSS fibres. Using this method, we find that the best predictor of gas-phase metallicity is Φ, in agreement with the results for stellar metallicity (which were derived using IFS).

This work is organized as follows: after introducing the data and the sample (Section 2), we present the results of our analysis (Section 3) and a discussion of the implications (Section 4); we conclude with a concise summary of our findings (Section 5). Throughout our work, we adopted a ΛCDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$.

2 DATA AND SAMPLE

We use publicly available data and measurements based on the SDSS Data Release 7 (SDSS DR7; Abazajian et al. 2009). Throughout this work, unless otherwise specified, we use spectroscopic redshifts (z), stellar masses (M_*), star formation rates (SFR), specific SFR (sSFR) and emission line fluxes obtained from the publicly available SDSS DR7 MPA/JHU catalogue (Kauffmann et al. 2003a; Brinchmann et al. 2004; Salim et al. 2007). The uncertainties on M_* and SFR were calculated as the semidifference between the 84th and 16th percentiles in the posterior distribution of each parameter. In addition, we use half-light radii (R_e) from the NYU Value Added Catalogue (NYUVAC; Blanton et al. 2005).

We assume uniform uncertainties on log R_e equal to 0.05 dex, determined by comparing log R_e to the corresponding values measured using the Multi Gaussian Expansion fitting technique (MGE; Emsellem, Monnet & Bacon 1994), as implemented by Cappellari (2002). 0.05 dex is obtained from rms/$\sqrt{2}$, where rms is the observed root mean square about the best-fitting linear relation between the NYUVAC and MGE values of log R_e.

The conversion between apparent and physical size is performed using the redshift-determined angular diameter distance in the adopted cosmology (Section 1). We assume that peculiar velocities are negligible, because our sample consists primarily of field and group galaxies: at the lowest redshift ($z = 0.01$), the error in distance due to a peculiar velocity of 300 km s^{-1} is $\approx$10 per cent or 0.04 dex, less than the measurement uncertainty on R_e.

For each galaxy, we use M_* as a proxy for the total mass M; we further define a proxy for the gravitational potential at one effective radius: $\Phi \equiv M_*/R_e$ and a proxy for the average surface mass density within one effective radius: $\Sigma \equiv M_*/R_e^2$. These definitions are only first-order approximations, that ignore variations both in the intrinsic shape and in the mass-to-light ratio of galaxies (e.g. Cappellari et al. 2013a). Nevertheless, it has been shown that the photometric proxies adopted here are precise enough for a comparative analysis of the stellar population properties, with results in excellent qualitative agreement with the equivalent spectroscopic proxies (i.e. $M_{spec} \equiv \sigma^2 R_e$, $\Phi_{spec} \equiv M_{spec}/R_e$, and $\Sigma_{spec} \equiv M_{spec}/R_e^2$; B18).

2.1 Sample selection

In order to guarantee that our results are representative, our main selection criteria are designed to build a volume-limited sample, with redshift $0.01 \leq z \leq 0.065$ and stellar masses $M_* > 10^9\,M_\odot$. Compared to the current literature, our selection encompasses a smaller and closer volume: other studies have used $0.005 < z < 0.25$ (T04) or $0.07 < z < 0.3$ (Telford et al. 2016). Given that our study focuses on the importance of fibre coverage, the redshift range is the most important selection criterion. However, our results are unchanged if we drop the volume-limited condition to adopt either of the two samples cited (Appendix A). The precise limits of the volume-limited sample do not affect our results: we repeated our test with $z < 0.08$ and $M_* > 10^{9.5}\,M_\odot$, and $z < 0.11$, and $M_* > 10^{10}\,M_\odot$, and found the results to be qualitatively unchanged.

Our quality cuts follow the criteria of Mannucci et al. (2010) and Telford et al. (2016): we require a signal-to-noise ratio (SNR) of at least 25 in the Hα line, of 5 in the Hβ line and of 3 in both the [S II]λ6717 and [S II]λ6731 lines (emission-line wavelengths are in Å unless otherwise specified). These criteria ensure that our sample is not biased in metallicity (see Telford et al. 2016, their fig. 1). We reject galaxies contaminated by active galactic nuclei (AGNs) emission using the classification of Kauffmann et al. (2003b), which relies on the BPT diagram (Baldwin, Phillips & Terlevich 1981). The precise threshold adopted to reject AGNs does not affect our results, as expected from other studies (Kewley & Ellison 2008). We further impose a lower limit on the equivalent width of the Hα emission line, EW(Hα), in order to remove galaxies where the emission lines arise from gas ionized by low-mass, evolved stars rather than OB stars associated with star formation (Stasińska et al. 2008). The specific value of the minimum EW(Hα) does not affect our analysis: even adopting the most conservative selection found in the literature, our results are qualitatively unchanged (Lacerda et al. 2018, EW(Hα) > 14 Å). However, because this strict selection removes approximately 25 per cent of our sample, we adopt the less conservative estimate of Stasińska et al. (2008), EW(Hα) > 3 Å, as also suggested by other studies (Davies et al. 2014, Cid Fernandes et al. 2011; 17 galaxies do not meet this threshold). Our final sample consists of 68 959 unique galaxies. This is what we call the 'parent sample'.

2.2 Metallicity measurements

The emission-line fluxes are dust-extinction corrected using Balmer decrement Hα/Hβ. We assume the Cardelli, Clayton & Mathis (1989) dust extinction law and case B recombination (Hα/Hβ = 2.86 for $T_e = 10^4$ K Osterbrock & Ferland 2006). We measured the metallicities using the new metallicity calibrations proposed by Dopita et al. (2016, hereafter D16). This calibration is based on a new grid of photoionization models from the MAPPINGS V code (Sutherland et al.in prep.). The combination of [N II] λ6584/[S II] $\lambda\lambda$6717,6731 and [N II] λ6584/Hα provides a metallicity diagnostic that is robust against dust reddening and ionization effects. Throughout the paper, the gas-phase metallicity refers always to the values calibrated following D16, unless otherwise specified.

In order to explore the effect of different metallicity calibrations, we repeated our analysis with three different calibrations: (a) the metallicity measurements from T04, based on the R23 line ratios (R23 $\equiv$ ([O II]λ3727 + [O III]$\lambda\lambda$4959, 5007)/Hβ) and publicly available through the MPA/JHU catalogue ((O/H)$_{\rm T04}$); (b) the metallicity calibrated using the O3N2 ratio ((O/H)$_{\rm O3N2}$; Marino et al. 2013); and (c) the metallicity calibrated using the N2 ratio ((O/H)$_{\rm N2}$; Marino et al. 2013). The different calibrations yield the same qualitative results, as reported in Section 3.6.

2.3 Sample properties

The properties of the sample are summarized in Fig. 1: the redshift-mass plane and the mass distribution (panels b and c) show that there is some incompleteness below $M_* = 10^{9.5}\,{\rm M}_\odot$, due to the SNR requirements. To address this issue, we use $V_{\rm max}$ weighting: the mass function is shown again in Fig. 2(a), where the dashed histogram is the mass function of our sample, the filled blue histogram is the $V_{\rm max}$-weighted mass function, and the dashed red line is the best-fitting Schechter function. We find $\log \mathcal{M}^*[{\rm M}_\odot] = 10.61$ and $\alpha = -1.58$ (where $\mathcal{M}^*$ and α are defined following Kelvin et al. 2014). Within the uncertainties, these values are consistent with the parameters of the Schechter function of disc-dominated galaxies (Kelvin et al. 2014, their table 3). The difference between the empty dashed and blue filled histograms highlights that the SNR requirements on the emission lines affect our sample selection, but thanks to our selection criteria, the correction to the mass function is small: repeating our analysis with uniform weighting leaves our results qualitatively unchanged.

Another possible bias that SNR cuts might introduce is against low sSFR galaxies: at fixed mass, galaxies with high sSFR are more likely to meet the quality cut than low sSFR galaxies. In order to assess the importance of this bias, we study the star-forming main sequence (SFMS; Noeske et al. 2007; Renzini & Peng 2015). We fit a straight line to the SFMS, using the least trimmed squares algorithm (Rousseeuw & Leroy 1987; Rousseeuw & Driessen 2006) in the free implementation `lts_linefit` (Cappellari et al. 2013a). We find a slope $a = 0.581 \pm 0.003$, with rms = 0.322 dex (Fig. 2b, solid red line). The general shape of the relation is correctly reproduced (Fig. 2b), however we notice that our best-fitting slope is flatter than the ridge line shown in Renzini & Peng (2015, dotted blue line in fig. 2b). This discrepancy is partly due to aperture bias at the high-mass end. Because of our redshift range, we probe only the innermost region of the highest mass galaxies (on average, the largest). For these galaxies, we may measure a lower-than-average SFR. However, this bias does not affect our results, because we concentrate here on aperture-matched sampling (see Section 2.4). For instance, by repeating the fit with the aperture-matched subsample with $R_{\rm fib}/R_{\rm e} = 1.5$ we find a steeper best-fitting slope of the SFMS $a = 0.62 \pm 0.01$.

We are mainly interested in the relation between three galaxy observables: mass, size, and metallicity; Fig.1 provides a summary of the distribution of these quantities for our sample. Notice the correlation between size and mass (panel e), the mass–metallicity relation (panel h), and the lack of correlation between metallicity and size (panel i). We use `lts_linefit` to fit a straight line to the mass–size relation (panel e) and find:

$$\log\ R_{\rm e}\,[{\rm kpc}] = (0.193 \pm 0.002)\log {\rm M}_*\,[{\rm M}_\odot] - (1.478 \pm 0.016) \quad (1)$$

with rms = 0.195; the best-fitting line is depicted as a solid green line in Fig.1(e), the dashed green lines are offset by $\pm$rms. Our results are in statistical agreement with the literature: Shen et al. (2003) studied the mass–size relation for a sample of SDSS late-type galaxies (defined as having Sérsic index $n < 2.5$; Sérsic 1968; Shen et al. 2003). They fit a relation of the form:

$$R_{\rm e}\,[{\rm kpc}] = \gamma \left(\frac{M}{{\rm M}_\odot}\right)^{\alpha} \left(1 + \frac{M}{M_0}\right)^{\beta-\alpha}. \quad (2)$$

We used an uncertainty weighted least-squares algorithm to fit this function to our data, and find $\alpha = 0.15 \pm 0.01$, $\beta = 0.57 \pm 0.28$, $\gamma = 0.09 \pm 0.01$, and $M_0 = (1.2 \pm 1.1) \times 10^{11}\,{\rm M}_\odot$, in agreement with the values in the literature ($\alpha = 0.14$, $\beta = 0.39$, $\gamma = 0.10$, and $M_0 = 3.98 \times 10^{10}\,{\rm M}_\odot$; Shen et al. 2003, their table 1). The best-fitting relation from Shen et al. (2003) is the dashed green line in Fig. 1(e); our best-fitting for the same function overlaps and is not depicted for clarity. We notice that the parameters β and M_0 in our fit are unconstrained, because of the large uncertainties on the best-fitting values: this fact is a consequence of the relative lack of massive galaxies in our sample compared to Shen et al. (2003), and emphasizes that a linear fit in log-space is adequate to model the mass–size relation for the range of masses and sizes considered in this work.

2.4 Aperture-matched subsamples

In order to establish the effect of metallicity gradients on the mass–metallicity relation, we divide the volume-limited sample in four subsets, based on the ratio between the *apparent* galaxy size and the fibre radius: $R_{\rm fib}/R_e = 0.5$, $R_{\rm fib}/R_e = 1$, $R_{\rm fib}/R_e = 1.5$ and $R_{\rm fib}/R_e = 2$. Our results hold for apertures up to $R_{\rm fib}/R_e \sim 2$, however beyond this limit the sample size is too small (<1000 galaxies) and our results are not significant. The smallest aperture is constrained by the requirement that we measure a representative fraction of the galaxy surface area. For an exponential light profile, an aperture of radius $0.5\,R_{\rm e}$ contains a fraction of the total light equal to 0.21: below this critical fraction the measured gas-phase metallicity does not necessarily represent the global metallicity of the galaxy (Kewley, Jansen & Geller 2005). In addition, for apertures smaller than $0.5\,R_{\rm e}$ our proxies for the potential Φ and surface mass density Σ become inadequate: for example, the relevant potential Φ cannot be estimated as $M_*/R_{\rm e}$, but we have to use the ratio between the stellar mass enclosed in the fibre ($M_{*,\,\rm fib}$) and the fibre radius ($R_{\rm fib}$, expressed in physical units). For these reasons, we do not consider apertures smaller than $0.5\,R_{\rm e}$.

The median fibre coverage in our sample is $R_{\rm fib}/R_{\rm e} = 0.5$ (Fig. 1u), with an extended tail to $R_{\rm fib}/R_{\rm e} = 2$. In order to increase the size of each subsample, we introduce a tolerance factor $t = 0.13$: for a

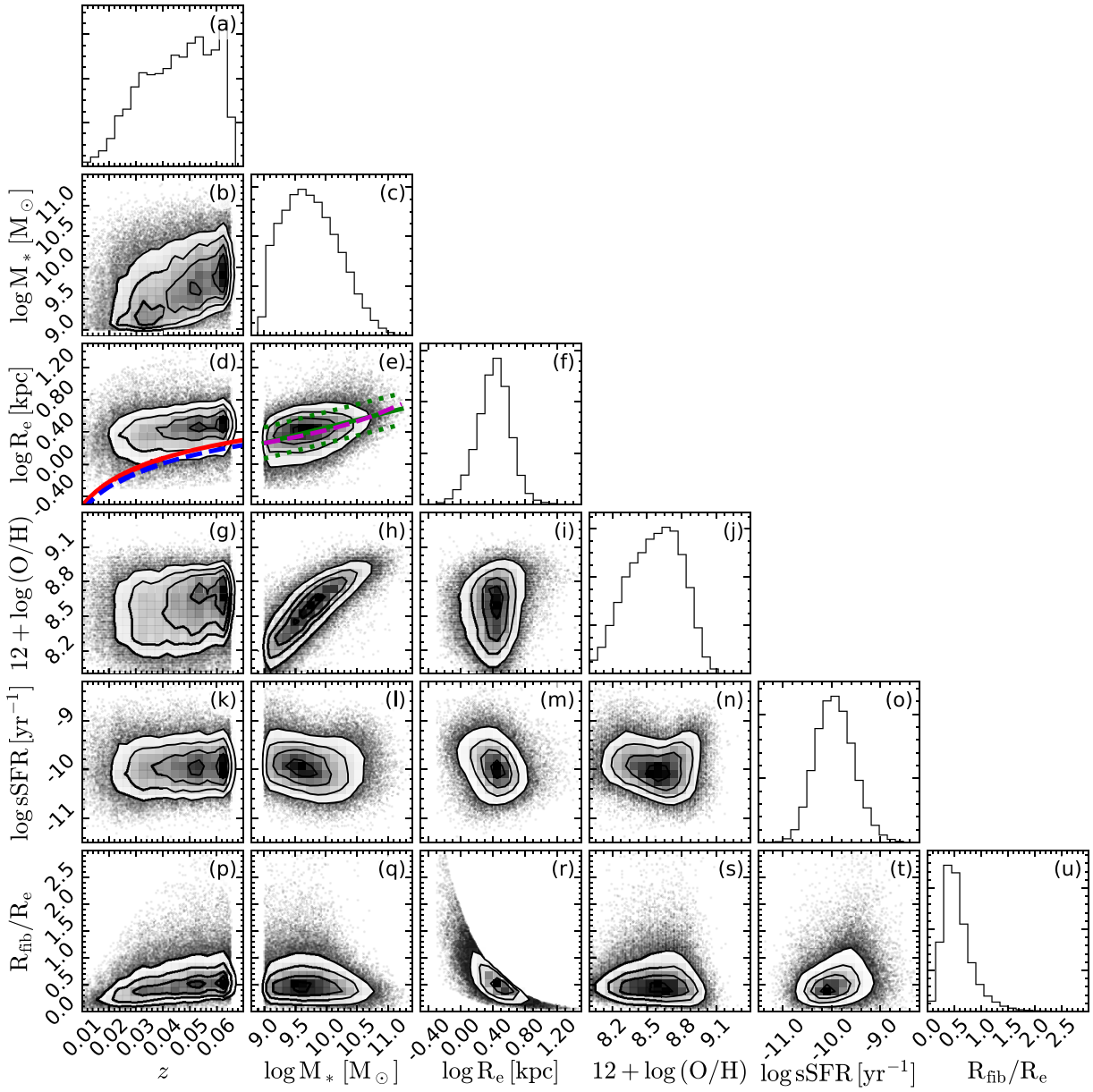

Figure 1. Summary of the physical properties of our sample. The solid black contours enclose the 11th, 39th, 68th, and 86th percentiles. The solid red line in panel d shows the radius of the SDSS fibre as a function of redshift, while the dashed blue line is the median full width half maximum of the r'-band point spread function. The mass–size relation is highlighted by the solid green line (panel e; the dotted green lines are offset by the root mean square). The dashed magenta line is the mass–size relation from Shen et al. (2003). There is a sharp contrast between the tight correlation observed in the mass–metallicity plane (panel h) and the lack of correlation between size and metallicity (panel i). Our sample has a median fibre coverage $R_{\rm fib}/R_{\rm e} = 0.5$ (panel u).

given ratio $R_{\rm fib}/R_{\rm e} = r$, we select the corresponding subsample using the inequalities $(1 - t)\, r < R_{\rm fib}/R_{\rm e} < (1 + t)\, r$. The value $t = 0.13$ guarantees that each of the aperture-matched samples contains the maximum number of galaxies, without overlapping. We repeated the analysis with $t = 0.10$ and $t = 0.05$ and find that the results do not depend on the choice of t. For reference, given the adopted uncertainty on log $R_{\rm e}$, the relative uncertainty on $R_{\rm e}$ is ≈0.12.

3 RESULTS

Before presenting our main results, we quantify the effect of aperture bias (Section 3.1). We then proceed to study the relation between stellar mass (M_*), size ($R_{\rm e}$), and gas-phase metallicity (O/H) by considering subsamples of fixed mass and size (Section 3.2). Next we show how O/H is distributed on the mass–size plane (Section 3.3). Finally, we present a quantitative study of the potential–metallicity relation (Φ–O/H) and show that it has less scatter and less residual trends with size than both the mass–metallicity (M–O/H) and surface density–metallicity relations (Σ–O/H; Section 3.4). We conclude by showing that our results do not arise due to the effects of sample or measurement bias (Section 3.5) and do not depend on the specific metallicity calibration adopted (Section 3.6).

3.1 Aperture bias

In order to study the correlation between gas-phase metallicity and the physical size of galaxies at fixed mass, it is important to assess and quantify the bias due to fixed aperture size. Past studies have focussed on the redshift-dependent effect of aperture bias: at fixed

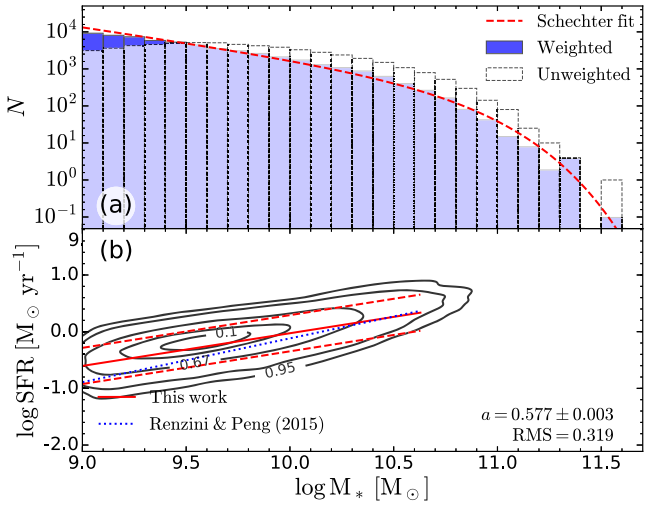

Figure 2. The mass function of the parent sample is compared to the mass function measured in the local Universe ($z < 0.065$; Kelvin et al. 2014, dashed red line; panel a). The star-forming main sequence (SFMS) of the parent sample is shown in panel b; the solid black contours encompass the 95th, 90th, 67th, 30th, and 10th percentiles of the data distribution. The best-fitting line has slope $a = 0.581 \pm 0.003$ (solid red line) and rms = 0.322 dex (dashed red lines). The approximate location of the ridge-line of the volume-weighted SFMS measured by Renzini & Peng (2015) is depicted by the dotted blue line.

Table 1. log O/H variation with size at fixed mass and with mass at fixed size.

Figure (1)	$\log R_e$ (kpc) (2)	$a_M \pm \sigma_a$ (3)	a_M/σ_a (4)	rms (5)	Δlog O/H (6)
(4a)	[0.52; 0.62]	0.45 ± 0.01	65.52	0.12	0.89
(4b)	[0.42; 0.52]	0.48 ± 0.01	82.83	0.12	1.15
(4c)	[0.32; 0.42]	0.52 ± 0.01	69.24	0.12	1.02
(4d)	[0.27; 0.37]	0.54 ± 0.01	54.32	0.13	0.99
Figure (1)	$\log M_*$ ($M_\odot$) (2)	$a_R \pm \sigma_a$ (3)	a_R/σ_a (4)	rms (5)	Δlog O/H (6)
(4e)	[10.4; 10.6]	-0.38 ± 0.04	10.66	0.10	-0.18
(4f)	[9.9; 10.1]	-0.51 ± 0.02	21.10	0.12	-0.33
(4g)	[9.4; 9.6]	-0.64 ± 0.02	31.11	0.13	-0.43
(4h)	[9.1; 9.3]	-0.61 ± 0.02	27.15	0.13	-0.45

Notes. For each of the panels in Fig. 4, we detail (1) the size range (panels a–d) or the mass range (panels e–f); (2) the best-fitting slope of the relation between log O/H and log M_* (panels a–d) or between log O/H and log R_e (panels e–f); (3) the statistical significance of the trend (in units of the standard deviation σ_a); (4) the observed root mean square around the best-fitting relation (5) and the total variation in log O/H over the observed range (6).

physical size a galaxy appears smaller with increasing distance, therefore the constant aperture of SDSS single-fibre spectroscopy probes increasing fractional areas with increasing redshift. This effect amounts to a maximum increase in radial coverage between ≈3 (Telford et al. 2016) and ≈40 (T04), depending on the redshift range selected. At fixed mass and redshift, the SDSS fibre covers a decreasing fraction of the galaxy light with increasing galaxy size: for the smallest galaxies the fibre encompasses all the light, whereas for the largest galaxies the fibre covers only the innermost, relatively metal rich regions. This observational effect induces an

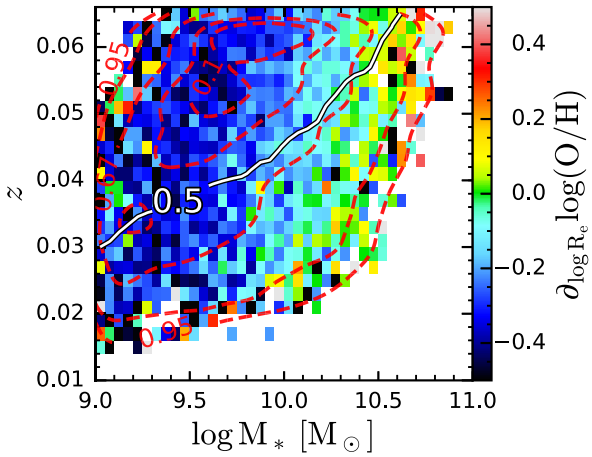

Figure 3. We find a negative metallicity gradient with galaxy size ($\partial_{\log R_e} \log (O/H) < 0$); given that the gradient due to aperture bias is positive, the negative gradients must be physical. For each bin at fixed mass and redshift, the colour map indicates the value of the slope of the best-fitting linear relation between metallicity and size ($\partial_{\log R_e} \log (O/H)$). The dashed red contour lines enclose the 95th, 90th, 67th, 30th, and 10th percentiles of the galaxy distribution, while the solid white curve is the locus where the median fibre coverage is $R_{fib}/R_e = 0.5$: below this line the SDSS fibres might not measure accurate global metallicities.

artificial correlation between size and metallicity at fixed mass and redshift.

However, for most of our sample, we find that size and metallicity are anticorrelated: in Fig. 3, we show how the size–metallicity correlation varies on the mass–redshift plane. For each bin in mass and redshift, we use `lts_linefit` to measure the best-fitting linear slope between log(O/H) and log R_e ($\partial_{\log R_e} \log (O/H)$, represented as a colour map in Fig. 3; we omit bins with less than 20 galaxies, which account for a fraction of the sample of less than 0.05). The solid white curve is the locus where the median fibre coverage is 0.5: below this curve the SDSS fibre coverage is insufficient to measure a representative O/H (see Section 2.4 and Kewley et al. 2005). Aperture bias will tend to produce a positive gradient with galaxy size, because for the largest galaxies the SDSS fibre probes only the innermost, most metal-rich region. Therefore, the negative gradients observed in Fig. 3 must be physical (and indeed must be more negative than observed if residual aperture bias exists). These negative gradients demonstrate the physical anticorrelation between size and metallicity, independent of galaxy mass.

3.2 Gas-phase metallicity at fixed mass and at fixed radius

We now study the relation between mass, size, and metallicity firstly by fixing size and letting mass vary, and then by fixing mass and letting size vary. In order to avoid aperture bias, we consider the aperture-matched subsample with $R_{fib}/R_e = 1$, which is equivalent to using IFS data and summing the spectra over an aperture of radius 1 R_e, an approach widely used in the literature (e.g. Scott et al. 2009; McDermid et al. 2015; Scott et al. 2017; B18; Li et al. 2018). We consider four samples at fixed size (a–d) and four samples at fixed mass (e–h); each sample has been centred along the best-fitting mass–size relation (Section 2.3). The width of the mass bins is 2 × 0.1 dex, or twice the measurement uncertainty on log M_*. The width of the size bins is 2 × 0.02 dex, inferred by substituting the width of the mass bins in the expression of the best-fitting mass–size relation (this is comparable to the estimated uncertainty on the size

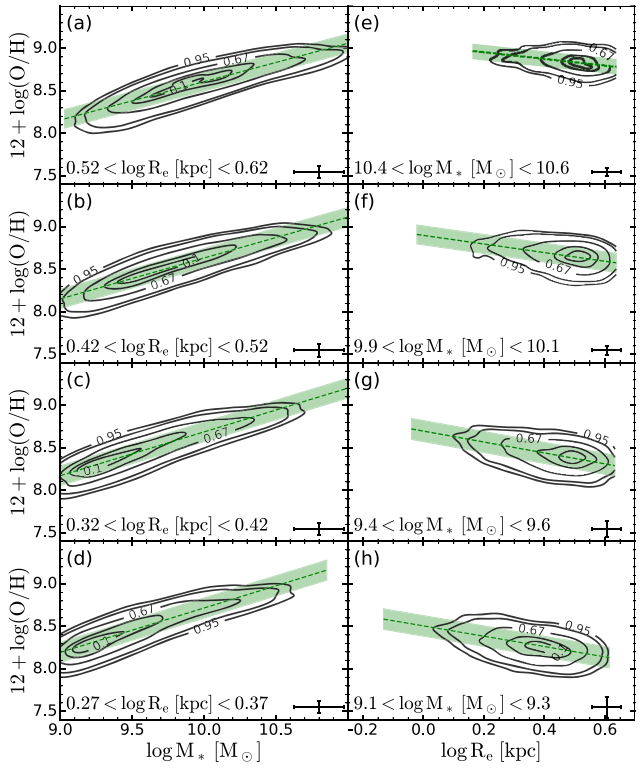

Figure 4. The dependence of gas-phase metallicity O/H on stellar mass M_* at fixed size R_e (panels a–d) and on R_e at fixed M_* (panels e–h). The bins of fixed mass and fixed size are selected along the mass–size relation. The contour lines enclose the 95th, 90th, 67th, 30th, and 10th percentiles of the data distribution, the dashed green lines are the best-fitting linear relations, and the green shaded regions span ±rms about the best fit. The clear anticorrelation between metallicity and galaxy size at fixed mass is physical, because aperture bias would create a positive correlation. The best-fitting values and rms for each panel are reported in Table 1, where we show that the strength of the mass–metallicity and size–metallicity correlations is comparable. The rms about the mass–metallicity relations is similar to the rms about the size–metallicity relations, indicating that mass and size are equally good predictors of metallicity.

measurements). The results are shown in Fig. 4: in the left column, we show the mass–metallicity relation for four bins at fixed physical size (panels a–d); in the right column, we show the size–metallicity relation for four bins at fixed mass (panels e–h). The mass or size range of each subsample is reported in the lower left corner of each panel, as well as in Table 1 (Column 1).

In each panel, the dashed green line represents the best-fitting linear relation and the green shaded region covers an offset of ±rms in metallicity. Although a linear relation does not describe accurately the mass–metallicity relation (the slope flattens at high masses), here our aim is to find the average gradient, and for this purpose a simple linear fit suffices. The best-fitting slope a and the rms about the best-fitting line are reported in Table 1 (Columns 3 and 5). We find everywhere a statistically significant trend (with significance $a/\sigma_a > 10$; Table 1 Column 4). The rms is very similar for the mass–metallicity and size–metallicity relations – so mass and size are equally good predictors of metallicity (which, given the scatter about the mass–size relation, would not be the case if one were a better predictor for metallicity than the other). The overall change in metallicity $\Delta\log$(O/H) is much smaller for the size–metallicity relation than for the mass–metallicity relation (Column 6), but this fact does not mean that size is less important than mass when it comes to metallicity, because the dynamic range in size is much smaller than the dynamic range in mass ($\Delta\log R_e \lesssim 0.6$ dex; panels e–h compared to $\Delta\log M_* \gtrsim 2$ dex; panels a–d). The difference in range between mass and size arises from three facts: in part, it reflects a real feature of the galaxies in our sample (Figs 1a and c: compare the width of the distributions of mass and size); in part, it results from the constraint that R_{fib}/R_e is within a narrow range of unity; and finally, it is part due to the flat slope of the mass–size relation ($a \approx 0.2$), so that we need to use a relatively wide range in mass to obtain the full range in size – which goes against the grain of this test (i.e. letting size vary at fixed mass).

A *caveat* of this analysis is that even the aperture-averaged metallicity within 1 R_e may still suffer from some degree of aperture bias, because earlier morphological types have more concentrated light profiles (Iglesias-Páramo et al. 2016). In order to measure an unbiased metallicity, Iglesias-Páramo et al. (2016) recommend using an aperture with radius equal to 2 R_e. Unfortunately our sample does not possess enough galaxies to repeat this quantitative analysis with the aperture-matched sample with $R_{fib}/R_e = 2$; however, we remark that our results are unchanged if we adopt the aperture-matched sample with $R_{fib}/R_e = 1.5$. In addition, if we use only galaxies with identical light profiles (selected by their measured Sérsic index), our results are unchanged (see Appendix B). We therefore conclude that aperture bias does not play a determining role in our results.

Despite these limitations, we find sufficient evidence to state that galaxy size affects the metallicity of the star-forming gas independent of galaxy mass; the logarithmic slope of the size–metallicity relation has opposite sign but comparable absolute value to the logarithmic slope of the mass-metallicity relation ($-1.2 < a_R/a_M < -0.8$; Table 1).

3.3 Gas-phase metallicity on the mass–size plane

In Fig. 5, we show the distribution of O/H on the mass–size plane (Cappellari et al. 2013b). Each row displays the results for a different sample: the parent sample (top row), and the aperture-matched subsamples with $R_{fib}/R_e = 0.5$ (second row), $R_{fib}/R_e = 1$ (third row), $R_{fib}/R_e = 1.5$ (fourth row), and $R_{fib}/R_e = 2$ (bottom row). The left column shows the distribution for the raw data (in square bins of 0.05 dex side, panels a–e); in the right column the data has been smoothed using the two-dimensional locally weighted robust regression technique (LOESS; Cleveland 1979; Cleveland & Devlin 1988), as implemented by Cappellari et al. (2013b); here, the circles represent individual galaxies, colour-coded by their inferred O/H. The solid lines in panel a are lines of constant median fibre coverage R_{fib}/R_e. The solid white curve indicates the locus where the (median) fibre coverage is $R_{fib}/R_e = 0.5$; this line is reproduced in all the panels. The region above this line consists of galaxies where the SDSS coverage is insufficient to measure a representative O/H (see Section 2.4 and Kewley et al. 2005). In all other panels, the dashed and the dotted black lines are lines of constant Φ and Σ, respectively.

For galaxies with $R_{fib}/R_e < 0.5$, we find that the lines of constant O/H are approximately vertical, i.e. lines of constant mass. However, for these galaxies the fractional fibre coverage is $R_{fib}/R_e \approx 0.25$, therefore: (i) the ratio M_*/R_e is not a good proxy for Φ and (ii) the measured metallicity is not representative of the galaxy metallicity. In fact, the observed trend above the white curve is exactly what is expected from aperture bias: at fixed M_*, the largest galaxies have the lowest fractional fibre coverage, therefore their fibre metallicity

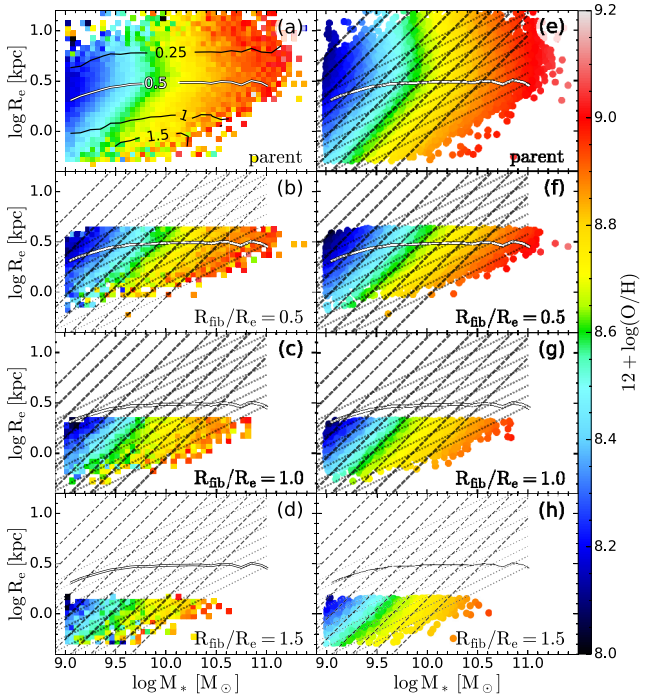

Figure 5. Gas-phase metallicity of the SDSS galaxies on the mass–size plane. The first row shows the results for the parent sample, the second, third, and fourth rows show the aperture-matched subsamples with $R_{fib}/R_e = 0.5$, $R_{fib}/R_e = 1$, and $R_{fib}/R_e = 1.5$. The panels in the left column show the raw (noisy) data, while the right-hand panels show the underlying distribution, reconstructed using the LOESS algorithm. In the top left panel, the solid lines are lines of constant (median) fibre coverage: the solid white curve with $R_{fib}/R_e = 0.5$ (replicated in all panels) is the critical line below which the SDSS fibres can accurately measure a representative O/H. Above this limit, the lines of constant metallicity are approximately vertical (panels a and f); however, in this region the galaxies have insufficient fibre coverage. Where the SDSS fibre covers at least $0.5R_e$ (i.e. below the white curve), metallicity is constant along lines of constant Φ (dashed black lines), rather than lines of constant mass (vertical) or constant Σ (dotted black lines; see panels b–e and f–j)

is higher than the average value for each galaxy. As a result, at fixed M_*, the largest galaxies are more metal rich than predicted by the local line of constant Φ.

In contrast, for galaxies with $R_{fib}/R_e > 0.5$ (i.e. below the solid white curve), we find that the lines of constant O/H are approximately aligned to lines of constant Φ. This is true for both the full sample (panels a and f) and for each of the individual aperture-matched samples (panels b–e and g–j). If we exclude the unreliable region with $R_{fib}/R_e < 0.5$, we find that the lines of constant metallicity (i) have approximately uniform slope across the full mass and size range, (ii) appear to be aligned along lines of constant Φ and (iii) nowhere in the valid part of the mass–size plane ($R_{fib}/R_e > 0.5$) are they aligned with lines of constant M or Σ.

We also observe that the zero-point of the metallicity relations within each aperture-matched subsample are different: at a given value of Φ, metallicity increases from the sample with $R_{fib}/R_e = 2$ to the sample with $R_{fib}/R_e = 0.5$. However, this change depends on the metallicity calibration adopted (Kewley & Ellison 2008), and may affect the relation for the parent sample (see Section 3.6).

3.4 The gas-phase metallicity–potential relation

We have seen that the gas-phase metallicity on the mass–size plane is approximately constant along lines of constant Φ (Section 3.3) rather than lines of constant M or Σ. We now seek to determine which of the three physical proxies M, Φ, and Σ is the best predictor of O/H. To this end, we consider the relation between O/H and each of M, Φ, and Σ, and we look at three metrics: (i) the rms about the running median, (ii) the Spearman rank correlation coefficient (ρ), and (iii) the presence and strength of trends between the residuals with respect to the median and the physical size (quantified using ρ_r, the relevant value of the Spearman rank correlation coefficient).

Firstly, we compare the gas-phase metallicity O/H with mass (M_*), gravitational potential (Φ), and surface mass density (Σ) for the full sample (Figs 6a–c). The galaxies are binned in both mass and metallicity, with each bin colour-coded by its median effective radius in physical units (R_e); the solid red line is the volume-weighted median metallicity in 30 equal-sized bins of the physical parameter, the dashed red lines are the 16th and 84th percentiles. The solid black lines are isodensity contours. In the top left corner of each panel, we report the rms, in the bottom left corner we report the Spearman rank correlation coefficient (ρ). The uncertainties on rms and ρ were estimated by bootstrapping 75 per cent of the sample 1000 times. All three panels show clear residual trends of O/H with R_e, which are visible as gradients in the colour hue in the main panels; in addition, these trends are quantified in the inset diagrams, where we show the residuals of each O/H relation as a function of log R_e, and where we report the value of ρ_r, the Spearman rank correlation coefficient for the distribution in the Δ log (O/H)–log R_e space.

For the parent sample (Figs 6a and b), the M–O/H and Φ–O/H relations are approximately equivalent, with similar values of the rms (0.135 versus 0.136) and ρ (0.819 versus 0.818), and similar (but opposite) ρ_r for the residuals as a function of log R_e (-0.281 versus 0.332). The Σ–O/H relation is clearly worse: it has larger rms (0.176), lower ρ (0.649), and higher ρ_r for the residuals as a function of log R_e (0.578).

In the third row of Fig. 6, we show the same diagram for the aperture-matched subsample with $R_{fib}/R_e = 1$ (panels g–i). The Φ–O/H relation is now better than both the Σ–O/H and M–O/H correlations: it has higher ρ, lower rms and it has no residual trends with galaxy size (ρ_r is -0.001 for the residuals of the Φ–O/H relation, compared to -0.221 and 0.345 for the residuals of the M–O/H and Σ–O/H relations). Equivalent results are seen also in the aperture-matched subsample with $R_{fib}/R_e = 0.5$ (panels d–f) and $R_{fib}/R_e = 1.5$ (panels j–l). For apertures larger than $R_{fib}/R_e = 1.5$, the relevant aperture-matched samples have too few galaxies to discriminate between the three metallicity relations. For apertures smaller than $\approx 0.5\,R_e$: (i) M_*/R_e might not be a good proxy for Φ in such small apertures and (ii) the metallicity measured within $R_{fib} < 0.5\,R_e$ is not representative of the total galaxy metallicity (see Section 3.3).

In conclusion, aperture-matched subsamples: (i) highlight the presence of residual size–metallicity trends at fixed mass and surface mass density and (ii) demonstrate that Φ is a slightly better (lower rms) and less biased (less residual with size) predictor of metallicity than either M or Σ.

3.5 Measurement bias and correlation between mass and size

One major possible source of bias is the correlation between stellar mass measurements and fibre size, because M_* is also estimated

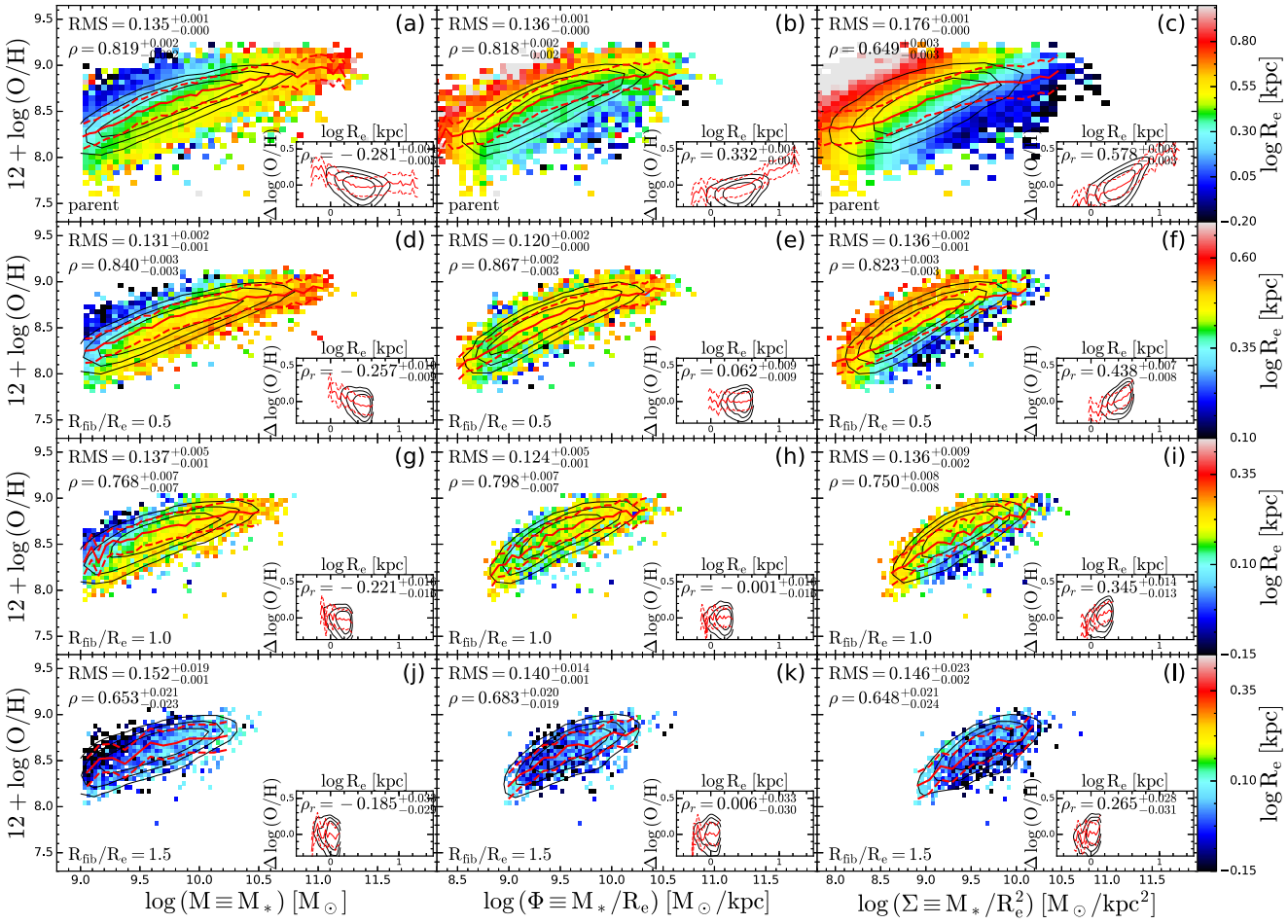

Figure 6. Gas-phase metallicity (expressed as 12 + log O/H) versus mass (M; left column), gravitational potential (Φ; central column), and surface mass density (Σ; right column). The first row (panels a–c) shows the results for the parent sample; the second, third, and fourth rows show the results for the aperture-matched samples with $R_{\rm fib}/R_{\rm e} = 0.5$, $R_{\rm fib}/R_{\rm e} = 1$, and $R_{\rm fib}/R_{\rm e} = 1.5$, respectively. The solid black contour lines enclose the 90th, 75th, and 50th percentiles of the galaxy distribution. The red solid line is the running volume-weighted median in 30 equal-sized bins of the physical parameter, the dashed lines are the 16th and 84th percentiles. In the inset diagram of each panel, we show the residuals with respect to the running median, as a function of log $R_{\rm e}$. The galaxies are colour-coded with their physical size ($R_{\rm e}$) to highlight the presence or absence of residual trends. For the aperture-matched samples, the Φ–O/H relations have less scatter (rms; top left corner of each panel), higher Spearman rank correlation coefficient (ρ; bottom left corner), and less residual trends with size (lower ρ_r in the inset diagrams) than both the M–O/H and Σ–O/H relations.

using fibre spectra and so these too vary with the coverage of the galaxy by the fibre. For this reason, we repeated our analysis using a purely photometric mass estimator that is completely independent of $R_{\rm fib}$ ($M_{*,\rm p}$; Taylor et al. 2011), we obtain the same results as using M_* $M_{*,\rm p}/R_{\rm e}$ is a better predictor of aperture-averaged metallicity than either $M_{*,\rm p}$ alone or $M_{*,\rm p}/R_{\rm e}^2$. We can therefore rule out the possibility that our results are caused by correlated errors between the mass measurements and the fibre coverage.

We can also exclude the possibility that our results arise from selection bias or correlated uncertainties. In principle, our SNR constraints (Section 2) might bias the sample against galaxies with low surface brightness, because a lower fraction of their light enters the SDSS fibre than for galaxies of equal total brightness but smaller apparent size. However, even by selecting our sample subject to $z < 0.065$ and $\Sigma_* \geq 10^{8.5}\,{\rm M}_\odot\,{\rm arcsec}^{-2}$ we find results consistent with those shown in Section 3.4: the M–O/H and Σ–O/H relations have clear residual trends with the physical size of galaxies, both for the full sample and for the aperture-matched subsamples. Hence, we rule out our results being a product of the selection criteria.

Correlated measurement uncertainties between M_* and $R_{\rm e}$ can artificially reduce the scatter in the Φ–O/H relation compared to the O/H–M relation. However, the measurements of M_* use a different measurement than the measurements of $R_{\rm e}$, which rules out correlated random uncertainties. Another possibility is the presence of correlated systematic errors due to fitting the galaxy light profiles with a pre-determined function (in our case, the family of Sérsic light profiles). However, our results are qualitatively unchanged if we replace the Sérsic $R_{\rm e}$ with the MGE $R_{\rm e}$, which are almost non-parametric (Emsellem et al. 1994; Cappellari 2002).

Finally, the physical correlation between M_* and $R_{\rm e}$ (Section 2) could artificially reduce the scatter in the Φ–O/H relation compared to the M–O/H relation. If we were using the wrong model to fit the empirical relations, we could be imposing a coherent structure in the residuals, which then combines with the correlation between M_* and $R_{\rm e}$ to induce (or remove) correlations between the residuals and galaxy size. However, this possibility is excluded because we use a non-parametric running median, which does not impose any pre-determined functional form on the metallicity relations and on the residuals.

It is worth remarking that – by construction – the observational uncertainty on M_*/R_e is larger than the observational uncertainty on M_* alone. For this reason, we conclude that whenever the observed rms about the Φ–O/H relation is equal to or smaller than the rms about the M–O/H relation, then the intrinsic scatter in the Φ–O/H relation is necessarily smaller than that in the M–O/H relation (we have ruled out the presence of correlated measurement uncertainties). The same argument holds for the Σ–O/H relations compared to both the M–O/H and Φ–O/H relations: the fact that the Σ–O/H relation has larger rms than both the M–O/H and Φ–O/H relations does not imply by itself that M or Σ are better predictors of O/H than Σ. However, the residuals of the Σ–O/H relation show a clear trend with galaxy size, whereas no such trend is observed for the residuals of the Φ–O/H relation. This fact suggests that at least part of the increase in rms from the Φ–O/H to the Σ–O/H relation is because Σ does not correctly take into account the metallicity information contained in the size of galaxies. We conclude therefore that, of the three structural parameters M ($\equiv M_*$), Φ ($\equiv M_*/R_e$), and Σ ($\equiv M_*/R_e^2$), the best predictor of O/H is Φ.

3.6 Alternative metallicity measurements

The results shown so far use the metallicity calibration of D16 (Section 2.2). However, adopting a different metallicity calibration affects the aperture-averaged metallicity measurements inside a constant aperture (Mannucci et al. 2010; S17). For this reason, it is important to test whether our results are altered when a different metallicity calibration is adopted. We repeated our full analysis using three alternative calibrations: $(O/H)_{T04}$, $(O/H)_{N2}$, and $(O/H)_{O3N2}$ (see again Section 2.2). In general, we find very good qualitative agreement between the different calibrations: the Φ–O/H relation has less observed scatter, higher ρ, and weaker residual trends with galaxy size than both the M–O/H and Σ–O/H relations, regardless of the specific metallicity calibration adopted.

The only major difference is the behaviour of the $(O/H)_{O3N2}$ distribution on the mass–size plane: in the top row of Fig. 7 (panels a and f), we find that the lines of constant metallicity (lines of approximately uniform colour hue) are not aligned with the lines of constant Φ (dashed black lines). In contrast to the results for the other calibrations, this observation is true both above and below the solid white curve, marking the locus of average fibre coverage equal to $R_{fib}/R_e = 0.5$. However, when we observe each of the aperture-matched subsamples, we find that the lines of constant metallicity are aligned with the lines of constant Φ, as observed for all the other calibrations. (panels b–e and g–j in Figs 5 and 7). The reason for this apparent discrepancy is that the zero-point differences of the aperture-matched subsamples for the Φ–O/H relations are much greater for the $(O/H)_{O3N2}$ metallicity than for any of the other calibrations: as a result, mixing together different aperture-matched samples in the parent sample introduces a stronger aperture bias in Fig. 7(f) than in Fig. 5(f). This emphasizes the importance of aperture-matched sampling.

We conclude that our main results are independent of the metallicity calibration adopted.

4 DISCUSSION

In the previous section, we have studied the correlation between the gas-phase metallicity (O/H) and the three photometric estimators for mass M, gravitational potential Φ, and surface mass density Σ. Here, we attempt to draw a consistent physical picture for the metallicity relations. We start by comparing our results to other studies of the global and local metallicity relations (Sections 4.1 and 4.2). We then compare the gas-phase metallicity trends to what has been observed for the stellar metallicity (Section 4.3). Finally, we present two physical interpretations in Section 4.4.

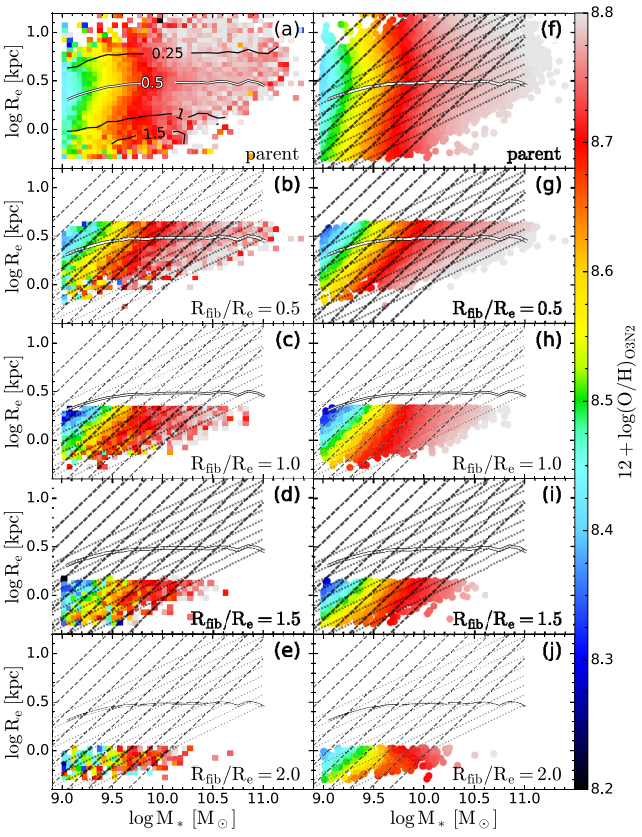

Figure 7. Gas-phase metallicity of the SDSS galaxies on the mass–size plane, using $(O/H)_{O3N2}$. This figure is the same as Fig. 5, except for the adopted metallicity calibration. In the parent sample, we find that $(O/H)_{O3N2}$ does not align with lines of constant Φ (black dashed lines; panels a and f). However, for each of the aperture-matched subsamples, the lines of constant $(O/H)_{O3N2}$ are always approximately aligned along lines of constant Φ (panels b–e and g–j). This behaviour is unique to the $(O/H)_{O3N2}$ calibration, because the difference in the zero-point of the Φ–O/H relations of each aperture-matched sample is highest for this calibration.

4.1 The aperture-averaged metallicity relations

In this work, we improved upon previous studies in two critical aspects. Firstly, by constructing a volume-limited sample, we ensure that our results are as representative as possible of the galaxy population in the local Universe (Section 2.3). Secondly, by studying the metallicity relations within aperture-matched samples, we minimize aperture bias.

The importance of a volume-limited approach is underlined by the controversy around the fundamental metallicity relation (Mannucci et al. 2010; Sánchez et al. 2013; Salim et al. 2014; Telford et al. 2016; S17; BB18). A specific limitation of studying the M–O/H relation in a volume-limited sample is that our results are not directly comparable to the current literature, which adopts different selection criteria (e.g. T04; Telford et al. 2016; S17). For instance, we find an rms of 0.135, larger than the values reported for both fibre-based studies (0.10 dex; e.g. T04) or IFU-based studies (0.06–0.10 dex; e.g. S17). However, this discrepancy is entirely explained

by the characteristics of our sample and by the metallicity calibration adopted. Our volume-limited sample contains a larger fraction of low-mass galaxies ($10^9 < M_* < 10^{10}\,M_\odot$) compared to both T04 and S17: the M–O/H relation has larger rms at the low-mass end than at the high-mass end (Guo et al. 2016, but see also Kewley & Elliso 2008, top panel of their fig. 2). For this reason, the observed rms depends primarily on the fraction of low-mass galaxies. For example, by adopting a volume-limited sample with $z \leq 0.08$ and $M_* > 10^{9.5}\,M_\odot$, we find rms = 0.11. When we repeat our analysis for the comparison samples, we find results that are consistent with the published values (Appendix A). As for the results obtained using IFU data (e.g. S17), the different mass function of our sample is compounded by the different quality of the data: synthetic aperture spectra constructed from IFU spectroscopy typically have much higher SNR than single-fibre spectroscopy. In addition, the metallicity calibration of D16 has larger scatter than other calibrations, as already demonstrated by other authors (S17; BB18). We conclude therefore that (i) for any given metallicity calibration, the average scatter about the M–O/H relation is larger than the typical values reported in the literature, which are based on magnitude-limited samples and (ii) given that the increased rms in the M–O/H relation is due to sample selection and metallicity calibration, it does not affect our comparative analysis between the three metallicity relations considered here. For this reason, we judge that the observed residual trend with size at fixed M reflects an underlying physical trend. Our conclusion is confirmed by other studies: the existence of a residual correlation between O/H and galaxy size at fixed mass has already been pointed out in previous works, both for the gas-phase metallicity (Ellison et al. 2008; Salim et al. 2014; Telford et al. 2016) as well as for the stellar metallicity (McDermid et al. 2015; Scott et al. 2017, B18; Li et al. 2018). Our results are in qualitative agreement with the current literature; however, we are the first to quantify the size–metallicity relation for the gas.

The importance of the aperture-matched samples is highlighted by the different behaviour of the metallicity distribution on the mass–size plane above and below the line of fibre coverage $R_{\rm fib}/R_{\rm e} = 0.5$. In the region above this line, metallicity appears to follow mass, whereas in the region below it metallicity follows potential. Given that this line also marks the region below which we can trust our metallicity and Φ measurements, we argue that aperture-matched sampling enables us to overcome aperture bias affecting previous fibre-based studies.

4.2 The local metallicity relations

Alongside the aperture-averaged M–O/H relation, several studies have found a correlation between the local gas-phase metallicity O/H($\boldsymbol{x}$) and a number of local properties of galaxies, typically defined on a physical scale of 1 kpc: local stellar mass surface density $\Sigma_*(\boldsymbol{x})$ (Rosales-Ortega et al. 2012; Sánchez et al. 2013; Barrera-Ballesteros et al. 2016), local escape velocity $V_{\rm esc(\boldsymbol{x})}$ (BB18) and local gas fraction $f_{\rm gas}(\boldsymbol{x}) \equiv \Sigma_{\rm gas}(\boldsymbol{x})/(\Sigma_{gas}(\boldsymbol{x}) + \Sigma_*(\boldsymbol{x}))$, with $\Sigma_{\rm gas}$ estimated from the Balmer decrement (BB18). In addition, if one uses the average metallicity and the effective radius of each galaxy as units of metallicity and radius, all galaxies have an average radial gradient of $\alpha_{\rm O/H} = -(0.10 \pm 0.01)$ [dex/$R_{\rm e}$] (radius–metallicity relation (R–O/H); Sánchez et al. 2014; Ho et al. 2015). Given that $f_{\rm gas}(\boldsymbol{x})$, $\Sigma_*(\boldsymbol{x})$, and $V_{\rm esc(\boldsymbol{x})}$ are correlated with one another, and that they can all be expressed as declining exponentials in radius, it is possible that one or more of the local relations are not independent.

Various authors have tried to explain which relations are physically motivated (e.g. Barrera-Ballesteros et al. 2016; BB18), however we call into question the explanations offered to date. Firstly, each of the three measurements comes with different random and systematic uncertainties, therefore a direct comparison of the scatter about the best-fitting metallicity relations cannot be used to assess which relation is intrinsically the tightest. We remark that even comparing the intrinsic scatter would be controversial, because such a comparison relies on the accurate estimate of the uncertainties. At this stage, the best observational tool to find the most fundamental relation is the study of residual trends. BB18 suggest that $f_{\rm gas}(\boldsymbol{x})$ is more important than $V_{\rm esc(\boldsymbol{x})}$ because the residuals about the best-fit $V_{\rm esc(\boldsymbol{x})}$–O/H($\boldsymbol{x}$) relation correlate strongly with $f_{\rm gas}(\boldsymbol{x})$, whereas the residuals about the $f_{\rm gas}(\boldsymbol{x})$–O/H($\boldsymbol{x}$) relation have a relatively weak correlation with $V_{\rm esc(\boldsymbol{x})}$. Although this argument is suggestive, there is no quantitative assessment of the relative strength of the residual correlations. In addition, (BB18) estimate $f_{\rm gas}(\boldsymbol{x})$ from the Balmer decrement (see Appendix C), which correlates observationally with metallicity via the dust fraction (Draine et al. 2014; Groves et al. 2015, see also Appendix C). With our fibre spectra, we cannot address these issues directly. However, for stellar metallicity, the local Mgb($\boldsymbol{x}$)–$V_{\rm esc(\boldsymbol{x})}$ relation is mirrored by an aperture-averaged Mgb–$V_{\rm esc}$ relation within one $R_{\rm e}$ (Scott et al. 2009). Arguably the same could be true for the gas metallicity.

4.3 Comparison to stellar metallicity

The results shown here for the gas-phase metallicity are in qualitative agreement with the trends of stellar metallicity ([Z/H]) for both early-type and late-type galaxies (ETGs and LTGs, respectively; Hubble 1936; Sandage 1961). For ETGs, the relation between [Z/H] and dynamical or stellar mass presents clear residual trends with galaxy size (McDermid et al. 2015; Scott et al. 2017) – in agreement with what we find here for O/H for star-forming galaxies (Section 3.4; Fig. 6a). B18 have shown quantitatively that for ETGs, [Z/H] correlates best with Φ rather than with M or Σ, regardless of the proxies (photometric or spectroscopic) used to estimate M, Φ, and Σ. [Z/H] appears to follow lines of constant σ (constant Φ in our terminology) also for LTGs (Li et al. 2018).

This qualitative agreement is particularly surprising for two reasons. Firstly, the underlying physics between stellar and gas-phase metallicity are different: stellar metallicity traces elements synthesized in all supernova types, whereas gas-phase metallicity (as measured here) traces only α-elements, produced only in Type-II supernovae. Moreover, the time-scale over which stellar and gas-phase metallicity evolve is also different: stellar metallicity measures the fraction of metals locked in the atmosphere of all stars, therefore it reflects the integrated star formation history of a galaxy; gas-phase metallicity measures instead the instantaneous abundance of metals in the star-forming gas, therefore it is influenced by gas inflows and by recent star formation (see e.g. González Delgado et al. 2014). With these caveats in mind, a qualitative comparison is still instructive.

A number of works have explored the dependence of local stellar metallicity on local galaxy properties: Emsellem et al. (1996) find a relation between the local value of the Lick index Mgb($\boldsymbol{x}$) and the local escape velocity ($V_{\rm esc(\boldsymbol{x})}$). Scott et al. (2009) observed that for the ETGs in the ATLAS^{3D} Survey, the local and aperture-averaged relations between Mgb and $V_{\rm esc}$ are consistent across all galaxies. If we interpret Mgb as a crude proxy for [Z/H], we expect a relation between the local stellar metallicity and $V_{\rm esc(\boldsymbol{x})}$.

B18 argue that the [Z/H]–Φ relation indicates that most stars in ETGs have formed *in situ,* which in this context means in a gravitational potential that is a monotonic function of the current

gravitational potential. For the stellar metallicity, this relation can be maintained even through dry mergers because the stars from accreted satellites are dispersed at a radius where the local potential is equal to the binding energy of the satellite (Villumsen 1983; McDermid et al. 2015). In a closed-box model, the *in situ* hypothesis predicts that the gas-phase metallicity follows the same relation as the stellar metallicity, as indeed we find in this study. Given that a closed-box model is not realistic, our results suggest that either outflows are not as important in determining the gas-phase metallicity, or they are regulated to a certain degree by the local gravitational potential.

4.4 Interpretation

The M–O/H relation is widely understood to be shaped by two physical processes: metal production and metal loss (e.g. T04; Lilly et al. 2013; Ho et al. 2015;BB18). To first order, we can assume that the mass of metals produced is proportional to the stellar mass, therefore the gas-phase metallicity O/H must be proportional to the ratio $\mathrm{O/H} \propto \mu^{-1} \equiv \Sigma_*(\boldsymbol{x})/\Sigma_{\rm gas}(\boldsymbol{x})$; in this framework, the local $\Sigma_*(\boldsymbol{x})$–$(\mathrm{O/H})(\boldsymbol{x})$ relation can be interpreted as an enrichment sequence (e.g. Barrera-Ballesteros et al. 2016). By integrating the $\Sigma_*(\boldsymbol{x})$–$(\mathrm{O/H})(\boldsymbol{x})$ relation (weighted with the light profile), one finds the aperture-averaged Σ_*–O/H relation. The second process, metal loss, depends on the value of $V_{\rm esc}$ (e.g. T04), which in turn is connected to the baryonic mass (estimated by M_*) by the baryonic Tully–Fisher relation (Bell & de Jong 2001). The M–O/H relation represents therefore a sequence of metal retention. In practice M_* and Σ_* are strongly correlated, therefore both the Σ–O/H and the M–O/H relations are jointly shaped by metal production and loss. Given the evidence from these past works, the correlation between O/H and Φ could arise from the combination of two effects: (i) metal loss is related to $M*$ (as argued by T04, thus explaining the M–O/H relation) and (ii) metal enrichment is driven by $f_{\rm gas}(\boldsymbol{x})$ (and therefore $\Sigma_*(\boldsymbol{x})$, as argued by Carton et al. 2015; Ho et al. 2015; Barrera-Ballesteros et al. 2016, thus explaining both the local $\Sigma_*(\boldsymbol{x})$–$(\mathrm{O/H})(\boldsymbol{x})$ relation and the aperture-averaged Σ–O/H relation). Each of these effects induces a correlation between O/H and $M_*/R_{\rm e}^{\rm n}$, with $n = 0$ and $n = 2$, respectively. Their combination produces an overall correlation with an intermediate value of n. In this framework, the R–O/H arises from the $\Sigma_*(\boldsymbol{x})$–$(\mathrm{O/H})(\boldsymbol{x})$ relation, and from the fact that Σ_* declines exponentially with radius for star-forming galaxies. The presence of residual trends with galaxy size in both the M–O/H and the Σ–O/H relations has a physical interpretation: at fixed Σ_*, larger galaxies have higher metallicity than smaller galaxies, because they are more massive and hence have lower escape fraction. Conversely, at fixed M, larger galaxies (which have lower Σ_*) have lower metallicity than smaller galaxies (which have higher Σ_*), because for a fixed escape fraction they produced less metals per unit mass.

Although this explanation is attractive, it presents two difficulties. Firstly, it requires some degree of fine tuning to explain why we find that the best correlation with $M_*/R_{\rm e}^{\rm n}$ has an exponent n that is very close to 1, the value corresponding to the proxy for the gravitational potential, Φ. Moreover, this hypothesis predicts that at the low-mass end of the relation O/H correlates best with M (because for low-mass galaxies the metallicity is dominated by outflows). Conversely, at the high-mass end, O/H should correlate best with Σ_*, because the escape fraction is small (the escape fraction is $f \lesssim 0.10$ for $M_* \gtrsim 3 \times 10^{10}\,\mathrm{M}_\odot$; Ma et al. 2016). Neither of these predicted trends is observed in our sample: the lines of constant gas-phase metallicity appear aligned along lines of constant Φ at every mass and size (Section 3.3).

Alternatively, the Φ–O/H relation might point to a direct physical link between the average depth of the gravitational potential Φ and the average metallicity O/H. Firstly, if we consider $M_*/R_{\rm e}$ as a crude proxy for the average $V_{\rm esc}$, the Φ–O/H relation reflects the observed link between the local metallicity $\mathrm{O/H}(\boldsymbol{x})$ and the local escape velocity $V_{\mathrm{esc}(\boldsymbol{x})}$. The existence of a link between $V_{\mathrm{esc}(\boldsymbol{x})}$ and $\mathrm{O/H}(\boldsymbol{x})$ (BB18) is in agreement with the observations for the stellar Mgb–$V_{\rm esc}$ relation (Emsellem et al. 1996; Scott et al. 2009). In addition, if one accepts that the Balmer decrement is a good estimator of $\Sigma_{\rm gas}$, we find that $\Sigma_{\rm gas}$ is correlated more tightly with Φ than with either Σ_* or M (Appendix C). Theoretically, the dynamical time of galaxies matches closely the time-scale of gas consumption (e.g. Lilly et al. 2013; Torrey et al. 2018), which suggests that the gravitational potential might play a key role in regulating star formation and therefore metallicity.

5 SUMMARY AND CONCLUSIONS

In this work, we examined the dependence of gas-phase metallicity (O/H) on the physical size of galaxies ($R_{\rm e}$) for a sample of $\sim 70\,000$ local star-forming galaxies drawn from the SDSS DR7. We use both a volume-limited parent sample, comprising galaxies with different radial coverage and three aperture-matched subsamples, selected to have a homogeneous fractional area coverage from the fixed-aperture SDSS fibres. For each sample, we compare O/H to three structural parameters: mass M (estimated by M_*), average gravitational potential Φ (estimated by $M_*/R_{\rm e}$) and surface mass density Σ (estimated by $M_*/R_{\rm e}^2$). Our results are as follows:

(i) We demonstrate the use of aperture-matched sampling based on single-fibre spectroscopy to study radially varying properties of galaxies, while minimizing aperture bias.

(ii) We show the existence of a size–metallicity relation at fixed mass, independent of aperture bias (Section 3.2; Fig. 4). The logarithmic slope of the size–metallicity relation is opposite in sign, and has comparable absolute value to the logarithmic slope of the mass–metallicity relation (Section 3.2; Table 1).

(iii) Over the region of the mass–size plane where we can reliably measure the metallicity, the lines of constant metallicity are closely aligned with the lines of constant Φ (Section 3.3; Fig. 5).

(iv) We find that – regardless of the sample selection adopted – the M–O/H and Φ–O/H relations have comparable root mean square (rms) and Spearman rank correlation coefficient (ρ; Figs 6a,b).

(v) For each of the aperture-matched subsamples, the Φ–O/H relation has the smallest rms, the highest ρ and the least-significant residual trends with galaxy size (Figs 6d–l).

(vi) We explore two possible explanations for the dependence of O/H on the physical parameters M, Φ, and Σ. The usual theory offered in the literature explains the Σ–O/H and M–O/H relations in terms of local enrichment (driven by $\Sigma_*(\boldsymbol{x})$) and global escape fractions (driven by M). We suggest, as an alternative, that both the M–O/H and Σ–O/H relations arise from a local relation between O/H and Φ. This hypothesis can be tested using integral field spectroscopy.

ACKNOWLEDGEMENTS

FDE acknowledges useful discussion with Prof. Roger L. Davies, Dr. Luca Cortese, Prof. Andrea Macciò, Dr. Chiaki Kobayashi, Dr. Philip Taylor, Dilyar Barat, and with the members of the SAMI

Galaxy Survey team. TMB is supported by an Australian Government Research Training Program Scholarship. Parts of this research were supported by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO; grant CE110001020), and the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D; grant CE170100013). BG gratefully acknowledges the support of the Australian Research Council as the recipient of a Future Fellowship (FT140101202).

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

This work made extensive use of the Debian GNU/Linux operative system, freely available at http://www.debian.org. We used the PYTHON programming language (van Rossum 1995), maintained and distributed by the PYTHON Software Foundation, and freely available at http://www.python.org. We acknowledge the use of SCIPY (Jones et al. 2001), MATPLOTLIB (Hunter 2007), EMCEE (Foreman-Mackey et al. 2013), ASTROPY (Astropy Collaboration et al. 2013), and PATHOS (McKerns et al. 2011). During the preliminary analysis we have made extensive use of TOPCAT (Taylor 2005).

REFERENCES

Abazajian K. et al., 2004, AJ, 128, 502
Abazajian K. N. et al., 2009, ApJS, 182, 543
Astropy Collaboration, 2013, A&A, 558, A33
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Barone T. M. et al., 2018, ApJ, 856, 64 (B18)
Barrera-Ballesteros J. K. et al., 2016, MNRAS, 463, 2513
Barrera-Ballesteros J. K. et al., 2018, ApJ, 852, 74 (BB18)
Bell E. F., de Jong R. S., 2001, ApJ, 550, 212
Blanton M. R. et al., 2005, AJ, 129, 2562
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bundy K. et al., 2015, ApJ, 798, 7
Cappellari M. et al., 2011, MNRAS, 413, 813
Cappellari M. et al., 2013a, MNRAS, 432, 1709
Cappellari M. et al., 2013b, MNRAS, 432, 1862
Cappellari M., 2002, MNRAS, 333, 400
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Carton D. et al., 2015, MNRAS, 451, 210
Cid Fernandes R., Stasińska G., Mateus A., Vale Asari N., 2011, MNRAS, 413, 1687
Cleveland W. S., 1979, J. Am. Stat. Assoc., 74, 829
Cleveland W. S., Devlin S. J., 1988, J. Am. Stat. Assoc., 83, 596
Croom S. M. et al., 2012, MNRAS, 421, 872
Davies R. L., Sadler E. M., Peletier R. F., 1993, MNRAS, 262, 650
Davies R. L., Kewley L. J., Ho I.-T., Dopita M. A., 2014, MNRAS, 444, 3961
de Zeeuw P. T. et al., 2002, MNRAS, 329, 513
Dopita M. A., Kewley L. J., Sutherland R. S., Nicholls D. C., 2016, Ap&SS, 361, 61 (D16)
Draine B. T. et al., 2014, ApJ, 780, 172
Ellison S. L., Patton D. R., Simard L., McConnachie A. W., 2008, AJ, 135, 1877
Emsellem E., Monnet G., Bacon R., 1994, A&A, 285, 723
Emsellem E., Bacon R., Monnet G., Poulain P., 1996, A&A, 312, 777
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Garnett D. R., Shields G. A., 1987, ApJ, 317, 82
González Delgado R. M. et al., 2014, ApJ, 791, L16
Groves B. A. et al., 2015, ApJ, 799, 96
Guo Y. et al., 2016, ApJ, 822, 103
Ho I.-T., Kudritzki R.-P., Kewley L. J., Zahid H. J., Dopita M. A., Bresolin F., Rupke D. S. N., 2015, MNRAS, 448, 2030
Hubble E. P., 1936, Realm of the Nebulae. Yale Univ. Press, New Haven
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Iglesias-Páramo J. et al., 2016, ApJ, 826, 71
Jones E. et al., 2001, SciPy: Open source scientific tools for Python, available at: http://www.scipy.org/
Jones D. H. et al., 2004, MNRAS, 355, 747
Kauffmann G. et al., 2003a, MNRAS, 341, 33
Kauffmann G. et al., 2003b, MNRAS, 346, 1055
Kelvin L. S. et al., 2014, MNRAS, 444, 1647
Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183
Kewley L. J., Jansen R. A., Geller M. J., 2005, PASP, 117, 227
Lacerda E. A. D. et al., 2018, MNRAS, 474, 3727
Lequeux J., Peimbert M., Rayo J. F., Serrano A., Torres-Peimbert S., 1979, A&A, 80, 155
Li H. et al., 2018, MNRAS, 476, 1765
Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119
Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140
McDermid R. M. et al., 2015, MNRAS, 448, 3484
McKerns M. M., Strand L., Sullivan T., Fang A., Aivazis M. A. G., 2011, in van der Walt S., Millman J., eds, Proceedings of the 10th Python in Science Conference. preprint (arXiv:1202.1056)
Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, MNRAS, 408, 2115
Marino R. A. et al., 2013, A&A, 559, A114
Moran S. M. et al., 2012, ApJ, 745, 66
Noeske K. G. et al., 2007, ApJ, 660, L43
Osterbrock D. E., Ferland G. J., 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei. University Science Books, Mill Valley, CA
Renzini A., Peng Y.-j., 2015, ApJ, 801, L29
Rosales-Ortega F. F., Sánchez S. F., Iglesias-Páramo J., Díaz A. I., Vílchez J. M., Bland-Hawthorn J., Husemann B., Mast D., 2012, ApJ, 756, L31
Rousseeuw P. J., Driessen K., 2006, Data Min. Knowl. Discovery, 12, 29
Rousseeuw P. J., Leroy A. M., 1987, Robust Regression and Outlier Detection. Wiley, New York
Salim S. et al., 2007, ApJS, 173, 267
Salim S., Lee J. C., Ly C., Brinchmann J., Davé R., Dickinson M., Salzer J. J., Charlot S., 2014, ApJ, 797, 126
Sánchez S. F. et al., 2012, A&A, 538, A8
Sánchez S. F. et al., 2013, A&A, 554, A58
Sánchez S. F. et al., 2014, A&A, 563, A49
Sánchez S. F. et al., 2017, MNRAS, 469, 2121
Sandage A., 1961, The Hubble Atlas of Galaxies. Carnegie Institution, Washington
Scott N. et al., 2009, MNRAS, 398, 1835

Scott N. et al., 2017, MNRAS, 472, 2833
Sérsic J. L., 1968, Atlas de Galaxias Australes. Observatorio Astronomico, Cordoba
Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978
Simard L., Mendel J. T., Patton D. R., Ellison S. L., McConnachie A. W., 2011, ApJS, 196, 11
Stasińska G. et al., 2008, MNRAS, 391, L29
Taylor E. N. et al., 2011, MNRAS, 418, 1587
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV. Astron. Soc. Pac., San Francisco, p. 29
Telford O. G., Dalcanton J. J., Skillman E. D., Conroy C., 2016, ApJ, 827, 35
Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, MNRAS, 404, 1775
Torrey P. et al., 2018, MNRAS, 477, L16
Tremonti C. A. et al., 2004, ApJ, 613, 898 (T04)
van Rossum G., 1995, Technical Report CS-R9526, Python tutorial, Centrum voor Wiskunde en Informatica (CWI), Amsterdam
Vila-Costas M. B., Edmunds M. G., 1992, MNRAS, 259, 121
Villumsen J. V., 1983, MNRAS, 204, 219
Worthey G., 1994, ApJS, 95, 107
Worthey G., Faber S. M., Gonzalez J. J., Burstein D., 1994, ApJS, 94, 687
York D. G. et al., 2000, AJ, 120, 1579
Zaritsky D., Kennicutt R. C., Jr, Huchra J. P., 1994, ApJ, 420, 87

APPENDIX A: NOTABLE SAMPLES

We favour a volume-limited sample to ensure that our conclusions are representative of the local Universe (Section 2). However, the results shown in Section 3 are largely independent of the sample selection criteria. Here, we repeat our analysis for two alternative samples selected following the criteria of two popular studies of the mass–metallicity relation: T04 and Telford et al. (2016).

T04 selected their sample imposing: (i) $0.005 < z < 0.25$, (ii) SNR(Hα) ≥ 5, SNR(Hβ) ≥ 5 and SNR(N IIλ6584) ≥ 5, (iii) uncertainty on z-band magnitude $\sigma(m_z) < 0.15$ mag, (iv) uncertainty on the Hδ_A index $\sigma(\mathrm{H}\delta_A) < 2.5$ Å, (v) uncertainty on the $D_n(4000)$ index $\sigma(D_n(4000)) < 0.1$, (vi) classified as star forming according to the BPT diagram of Kauffmann et al. (2003b), and (vii) uncertainty on stellar mass and metallicity less than 0.2 dex. Here, we reproduce their selection, but with two differences: we use data from SDSS DR7 instead of SDSS DR4, and we use the metallicity calibration from Dopita et al. (2016, see Section 2.2). Despite the different selection criteria between this sample and our sample (Section 2), the results of the comparative analysis are the same. In Fig. A1, we compare O/H to M, Φ, and Σ for the full sample (top row, panels a–c) and for the aperture-matched subsample with $R_{\rm fib}/R_{\rm e} = 1$ (bottom row, panels d–f). For the full sample, the M–O/H and Φ–O/H relations are equivalent: they have comparable rms and ρ and the residuals have similar but opposite correlations with galaxy size. However, for the aperture-matched samples, the best predictor of O/H is Φ, as we have found in Section 3 (Figs 6d–f).

Telford et al. (2016) selected their sample to minimize the bias against low-metallicity galaxies. Their constraints are as follows: (i) $0.07 < z < 0.30$, $A_V < 2.5$, (ii) Balmer decrement $F(\mathrm{H}\alpha)/F(\mathrm{H}\beta) > 2.5$, (iii) SNR(H$\alpha$) ≥ 25, SNR(Hβ) ≥ 5, SNR(S IIλ6717) ≥ 3, and SNR(S IIλ6731) ≥ 3, (iv) classified as star forming according to the BPT diagram of Kauffmann et al. (2003b), and (v) having valid size measurements in the catalogue of Simard et al. (2011). These selection criteria are identical to ours (Section 2.1), except for the redshift range: the volume considered by Telford et al. (2016) is ≈80 times larger than ours, but the low mass range is incomplete and would require large corrections, which we chose to avoid in this work.

Despite the different mass and redshift range, and without applying any correction for incompleteness, using this sample leads to the same conclusions as our volume-limited approach: the outcome of the comparative analysis between the metallicity relations with M, Φ, and Σ is the same as for our sample (Fig. A2). For the full sample, the M–O/H relation appears to be the best predictor of metallicity: it has lower rms and higher ρ than both the Φ–O/H and Σ–O/H relations (panels a–c). Even though the residuals of the M–O/H relation are correlated with $\log R_{\rm e}$ ($\rho_r = -0.184$; inset panel in Fig. A2a), this residual correlation is weaker than the residual correlations for the Φ–O/H and Σ–O/H relations (inset panels in Figs A2b and c).

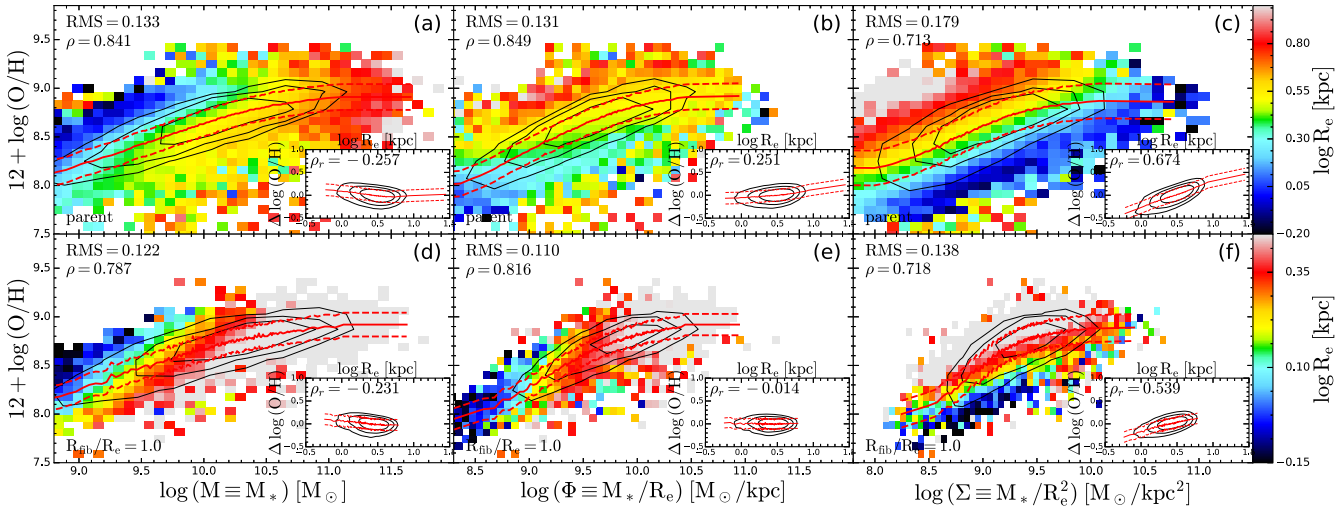

Figure A1. The same as Fig. 6, but for a sample reproducing the selection of T04. The top row shows gas-phase metallicity as a function of M, Φ, and Σ for the parent sample (panels a–c). The bottom row shows the results for the aperture-matched subsample with $R_{\rm fib}/R_{\rm e} = 1$ (d–f). For the parent sample, the best predictor of O/H is M (it has less scatter, higher ρ, and less residual trends with size than the O/H–Φ and O/H–Σ relations). For the aperture-matched subsample, the best predictor of O/H is Φ.

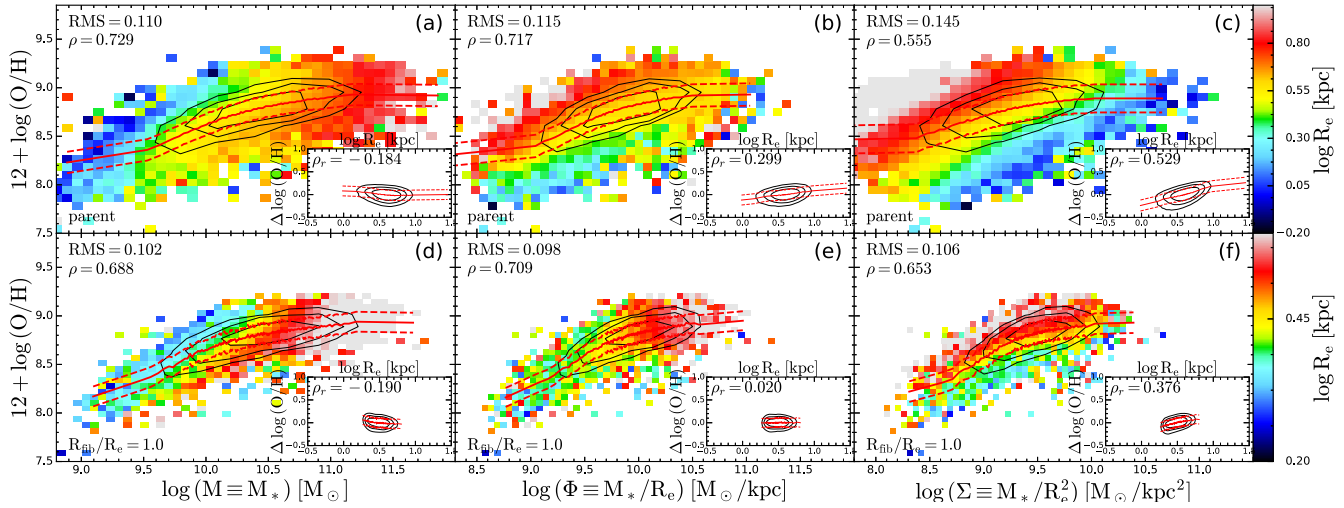

Figure A2. The same as Fig. 6, but for a sample reproducing the selection of Telford et al. (2016). The top row shows gas-phase metallicity as a function of M, Φ, and Σ for the parent sample (panels a–c). The bottom row shows the results for the aperture-matched subsample with $R_{\rm fib}/R_{\rm e} = 1$ (d–f). For the parent sample, the best predictor of O/H is M (it has less scatter, higher ρ, and less residual trends with size than the O/H–Φ and O/H–Σ relations). For the aperture-matched subsample, the best predictor of O/H is Φ.

However, when we study the relation for the aperture-matched subsample with $R_{\rm fib}/R_{\rm e} = 1$, we find that it is the Φ–O/H relation which is the best predictor of O/H: it has the lowest rms and the highest ρ. Moreover, the residuals of the Φ–O/H relation show no correlation with log$R_{\rm e}$ ($\rho = 0.020$; inset panel in Fig. A2e), unlike the residuals of the M–O/H and Σ–O/H relations ($\rho = -0.190$; inset panel in Fig. A2d and $\rho = 0.376$; inset panel in Fig. A2f).

This analysis reinforces our conclusion that our results do not depend on the sample selection criteria and that aperture-matched sampling is critical to recover an unbiased relation (i.e. aperture bias is more important than how the mass function is sampled).

APPENDIX B: SYSTEMATIC VARIATIONS IN THE LIGHT PROFILE

One major concern of our analysis is that studies stemming from IFU spectroscopy recommend using aperture-averaged metallicities within $2\,R_{\rm e}$ (Iglesias-Páramo et al. 2016), but here the corresponding aperture-matched sample (i.e. $R_{\rm fib}/R_{\rm e} = 2$) does not

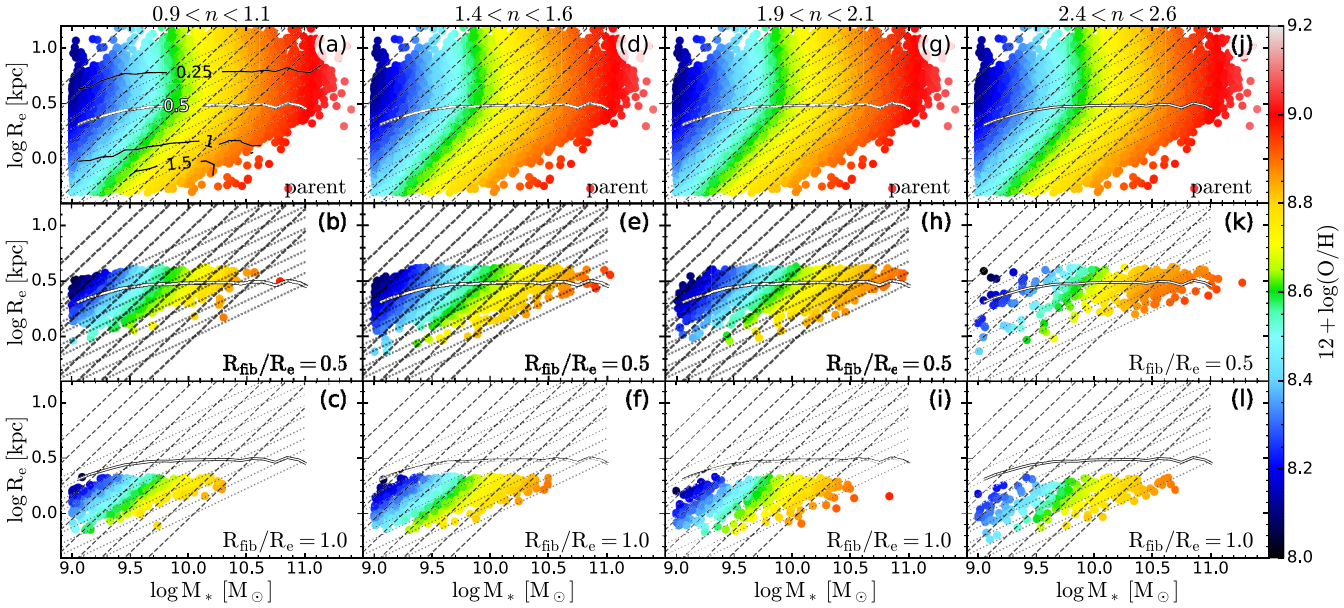

Figure B1. Distribution of the gas-phase metallicity O/H on the mass–size plane, divided by the best-fitting value of the *r*-band Sérsic index *n*. Each column shows the distribution for a different bin in *n*: $0.9 < n < 1.1$ (panels a–c), $1.4 < n < 1.6$ (panels d–f), $1.9 < n < 2.6$ (panels g–i), $2.4 < n < 2.6$ (panels j–l). Each row shows the distribution for a different aperture: the parent sample (top row, panels a, d, g, and j), the aperture-matched sample with $R_{\rm fib}/R_{\rm e} = 0.5$ (middle row, panels b, e, h, and k) and the aperture-matched sample with $R_{\rm fib}/R_{\rm e} = 1$ (bottom row, panels c, f, i, and l). In the aperture-matched samples, O/H is constant (i.e. approximately uniform colour hue) along lines of constant Φ (dashed black lines). This fact does not depend on the functional shape of the stellar light profile.

possess enough galaxies to perform a conclusive quantitative analysis. Iglesias-Páramo et al. (2016) argue that the systematic variation in the light profile among different spiral types introduces a morphology-dependent aperture bias. Early spirals (Sa) have by definition more prominent bulges than late spirals (Sc, Sd; Hubble 1936; Sandage 1961). As a result, the former have more concentrated light profiles than the latter, which in turn means that a smaller fraction of the disc light is included within $1\,R_e$ than for later-type galaxies.

In order to assess whether this bias is affecting our results, we used three methods. Firstly, we have shown qualitatively that even for the aperture-matched sample with $R_{fib}/R_e = 2$, O/H closely follows lines of constant Φ (Fig. 5j). Secondly, we show that our quantitative results hold for the sample with $R_{fib}/R_e = 1.5$ (panels j–l in Fig. 6). Here, we show that even by selecting galaxies with the same light profiles, our results are unchanged.

If the Φ–O/H relation were due to light profile dependent aperture bias still being present in our aperture-matched samples, we would expect that dividing these samples by the shape of the light profile showed a different dependence on mass and size. In Fig.B1, we reproduce the metallicity variation on the mass–size plane reconstructed with LOESS, divided in four narrow bins of the Sérsic index n: $0.9 < n < 1.1$ (first column), $1.4 < n < 1.6$ (second column), $1.9 < n < 2.1$ (third column), and $2.4 < n < 2.6$ (last column; for higher values of n our aperture-matched samples do not possess sufficient galaxies). Comparing Fig. B1 with Fig. 5 we find no qualitative difference: galaxies with different light profiles behave in the same way on the mass–size plane, so the fact that the lines of constant O/H are parallel to the lines of constant Φ is not an artefact of aperture bias.

APPENDIX C: USING THE BALMER DECREMENT TO ESTIMATE THE GAS SURFACE MASS DENSITY

BB18 use the Balmer decrement to estimate the gas surface mass density, by defining $\Sigma_{gas} \equiv 30\,A_V\,M_\odot\,pc^{-2}$, where A_V is the optical extinction. This estimator assumes that the dust-to-gas ratio ($\mathcal{D}$) is uniform across all galaxies; however, $\mathcal{D}$ depends on metallicity (Draine et al. 2014; Groves et al. 2015). For this reason, one should be careful when using this method to study the correlation between metallicity and the surface mass density of gas.

In Fig. C1, we show how Σ_{gas} (that is, A_V) varies across the mass-size plane. The solid lines are the loci of constant fibre coverage: our measurements of metallicity can be trusted only below the solid white curve (i.e. for $R_{fib}/R_e > 0.5$). The dashed (dotted) black lines are lines of constant Φ (Σ): it is clear that Σ_{gas} correlates with Φ, rather than with Σ. When we repeat the quantitative analysis of Fig. 6 by substituting Σ_{gas} in place of O/H, we find that the Σ_{gas}–Φ and Σ_{gas}–Σ relations have comparable rms, but the Σ_{gas}–Φ relation is statistically consistent with no residual trends with galaxy size, unlike the Σ_{gas}–Σ relation. We believe that this result is a consequence of the correlation between Φ and O/H, and between O/H and the Balmer decrement. If however we were to treat the Balmer decrement as an unbiased estimator of Σ_{gas}, the fact that Φ is the best predictor of Σ_{gas} could be used to explain the Φ–O/H relation in terms of effective yield and gas fractions (e.g. BB18). The fact that Σ_{gas} correlates better with Φ than with M or Σ raises the possibility that the Φ–O/H relation arises from the correlation between Φ and A_V and from A_V biasing the metallicity estimator O/H. However, the metallicity calibration adopted here uses line ratios that are sufficiently close in wavelength to be insensitive to the extinction correction (see Section 2.2 and Dopita et al. 2016). For this reason, we believe that O/H links Φ and A_V, rather than A_V linking Φ and O/H.

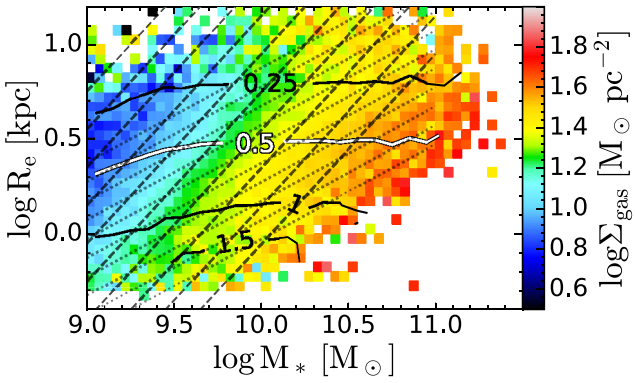

Figure C1. Gas surface density (Σ_{gas}, derived from the Balmer decrement) on the mass–size plane. The solid lines are lines of constant (median) fibre coverage $\kappa = R_{fib}/R_e$. Our fibre-based measurements of metallicity can be trusted only in the region below the solid white curve ($\kappa = 0.5$). The lines of constant Σ_{gas} (i.e. lines of uniform colour) are aligned to the lines of constant potential Φ (dashed black lines) rather than the lines of constant surface density Σ (dotted black lines).

This paper has been typeset from a TEX/LATEX file prepared by the author.